



I hereby certify that this correspondence is being deposited with the U.S. Postal Service with sufficient postage as First Class Mail, in an envelope addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on the date shown below.

Dated: January 19, 2004

Signature: 
(Thomas A. Miller)

Docket No.: 28944/39534
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Fathi Debili

Application No.: 10/621548

Confirmation No.: 01326

Filed: July 17, 2003

Art Unit: 1724

For: PROCESS, COMPUTERIZED DEVICE, AND
COMPUTER PROGRAM FOR ASSISTING
THE VOWELIZATION OF ARABIC
LANGUAGE WORDS

Examiner: Not Yet Assigned

CLAIM FOR PRIORITY AND SUBMISSION OF DOCUMENTS

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

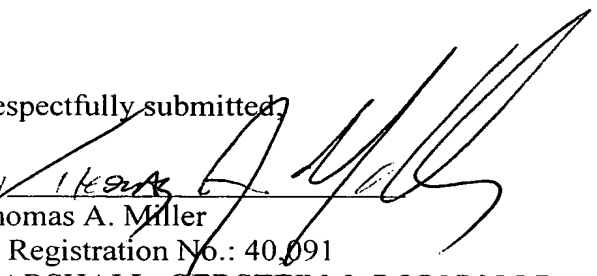
Applicant hereby claims priority under 35 U.S.C. 119 based on the following prior foreign application filed in the following foreign country on the date indicated:

Country	Application No.	Date
France	FR03 07665	June 25, 2003

In support of this claim, a certified copy of the said original foreign application is filed herewith.

Dated: January 19, 2004

Respectfully submitted,

By 
Thomas A. Miller

Registration No.: 40,091
MARSHALL, GERSTEIN & BORUN LLP
233 S. Wacker Drive, Suite 6300
Sears Tower
Chicago, Illinois 60606-6357
(312) 474-6300
Attorney for Applicant



BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 08 AOUT 2003

Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

Martine PLANCHE

INSTITUT
NATIONAL DE
LA PROPRIÉTÉ
INDUSTRIELLE

SIEGE
26 bis, rue de Saint Petersburg
75800 PARIS cedex 08
Téléphone : 33 (0)1 53 04 53 04
Télécopie : 33 (0)1 53 04 45 23
www.inpi.fr





26 bis, rue de Saint Pétersbourg
75800 Paris Cedex 08
Téléphone : 33 (1) 53 04 53 04 Télécopie : 33 (1) 42 94 86 54

BREVET D'INVENTION CERTIFICAT D'UTILITÉ

Code de la propriété intellectuelle - Livre VI



N° 11354*03

REQUÊTE EN DÉLIVRANCE

page 1/2

BR1

Cet imprimé est à remplir lisiblement à l'encre noire

DB 540 p. 1 / 210502

REMISE DE 25 COPIES DATE 25 JUIN 2003 LIEU 75 INPI PARIS N° D'ENREGISTREMENT 0307665 NATIONAL ATTRIBUÉ PAR L'INPI DATE DE DÉPÔT ATTRIBUÉE PAR L'INPI 25 JUIN 2003 Vos références pour ce dossier 030245 (facultatif)		1 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE CABINET PLASSERAUD 84, rue d'Amsterdam 75440 PARIS CEDEX 09	
Confirmation d'un dépôt par télécopie		<input type="checkbox"/> N° attribué par l'INPI à la télécopie	
2 NATURE DE LA DEMANDE		Cochez l'une des 4 cases suivantes	
Demande de brevet		<input checked="" type="checkbox"/>	
Demande de certificat d'utilité		<input type="checkbox"/>	
Demande divisionnaire		<input type="checkbox"/>	
Demande de brevet initiale ou demande de certificat d'utilité initiale		N° _____ N° _____	Date _____ Date _____
Transformation d'une demande de brevet européen Demande de brevet initiale		<input type="checkbox"/> N° _____	Date _____
3 TITRE DE L'INVENTION (200 caractères ou espaces maximum) Procédé, dispositif informatique et programme d'ordinateur pour l'aide à la voyellation de mots en langue arabe.			
4 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE FRANÇAISE		Pays ou organisation _____ N° _____ Date _____ Pays ou organisation _____ N° _____ Date _____ Pays ou organisation _____ N° _____ <input type="checkbox"/> S'il y a d'autres priorités, cochez la case et utilisez l'imprimé «Suite»	
5 DEMANDEUR (Cochez l'une des 2 cases)		<input checked="" type="checkbox"/> Personne morale <input type="checkbox"/> Personne physique	
Nom ou dénomination sociale		CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE - CNRS -	
Prénoms			
Forme juridique		Etablissement Public, Scientifique et Technologique EPST	
N° SIREN			
Code APE-NAF			
Domicile ou siège	Rue	3, rue Michel Ange 75016 PARIS Cédex 16	
	Code postal et ville		
	Pays	FRANCE	
Nationalité		Française	
N° de téléphone (facultatif)		N° de télécopie (facultatif)	
Adresse électronique (facultatif)			
<input checked="" type="checkbox"/> S'il y a plus d'un demandeur, cochez la case et utilisez l'imprimé «Suite»			

Remplir impérativement la 2^{ème} page



BREVET D'INVENTION CERTIFICAT D'UTILITÉ

REQUÊTE EN DÉLIVRANCE
page 2/2

BR2

REMISE DES PIÈCES DATE 25 JUIN 2003 LIEU 75 INPI PARIS N° D'ENREGISTREMENT NATIONAL ATTRIBUÉ PAR L'INPI 0307665 REF030245		Réservé à l'INPI DB 540 W / 210502
6 MANDATAIRE (s'il y a lieu)		
Nom		
Prénom		
Cabinet ou Société		Cabinet PLASSERAUD
N° de pouvoir permanent et/ou de lien contractuel		
Adresse	Rue	84, rue d'Amsterdam
	Code postal et ville	
	Pays	75009 PARIS
N° de téléphone (facultatif)		
N° de télécopie (facultatif)		
Adresse électronique (facultatif)		
7 INVENTEUR (S) Les inventeurs sont nécessairement des personnes physiques		
Les demandeurs et les inventeurs sont les mêmes personnes		<input type="checkbox"/> Oui <input checked="" type="checkbox"/> Non : Dans ce cas remplir le formulaire de Désignation d'inventeur(s)
8 RAPPORT DE RECHERCHE Uniquement pour une demande de brevet (y compris division et transformation)		
Établissement immédiat ou établissement différé		<input checked="" type="checkbox"/> Oui <input type="checkbox"/> Non
Paiement échelonné de la redevance (en deux versements)		Uniquement pour les personnes physiques effectuant elles-mêmes leur propre dépôt <input type="checkbox"/> Oui <input type="checkbox"/> Non
9 RÉDUCTION DU TAUX DES REDEVANCES Uniquement pour les personnes physiques		
		<input type="checkbox"/> Requise pour la première fois pour cette invention (joindre un avis de non-imposition) <input type="checkbox"/> Obtenue antérieurement à ce dépôt pour cette invention (joindre une copie de la décision d'admission à l'assistance gratuite ou indiquer sa référence) : AG
10 SÉQUENCES DE NUCLEOTIDES ET/OU D'ACIDES AMINÉS		
Le support électronique de données est joint		<input type="checkbox"/>
La déclaration de conformité de la liste de séquences sur support papier avec le support électronique de données est jointe		<input type="checkbox"/>
Si vous avez utilisé l'imprimé « Suite », indiquez le nombre de pages jointes		
11 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE (Nom et qualité du signataire) Eric BURBAUD 94-0304		VISA DE LA PRÉFECTURE OU DE L'INPI



26 bis, rue de Saint Pétersbourg
75800 Paris Cedex 08
Téléphone : 33 (1) 53 04 53 04 Télécopie : 33 (1) 42 94 86 54

BREVET D'INVENTION CERTIFICAT D'UTILITÉ

Code de la propriété intellectuelle - Livre VI



N° 11354*03

REQUÊTE EN DÉLIVRANCE

Page suite N° 1. / 1.

BR/SUITE

REMISE DES PIÈCES DATE 25 JUIN 2003 LIEU 75 INPI PARIS N° D'ENREGISTREMENT 0307665 NATIONAL ATTRIBUÉ PAR L'INPI		Réservé à l'INPI
Vos références pour ce dossier (facultatif) BFF03 0245		
4 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE FRANÇAISE		Pays ou organisation _____ N° _____ Date _____ Pays ou organisation _____ N° _____ Date _____ Pays ou organisation _____ N° _____ Date _____
5 DEMANDEUR (Cochez l'une des 2 cases)		<input checked="" type="checkbox"/> Personne morale <input type="checkbox"/> Personne physique
Nom ou dénomination sociale		ECOLE NORMALE SUPERIEURE LETTRE ET SCIENCES HUMAINES
Prénoms		
Forme juridique		
N° SIREN		
Code APE-NAF		
Domicile ou siège	Rue	15, Parvis René Descartes BP 700 69342 LYON CEDEX 07
	Code postal et ville	
	Pays	FRANCE Française
Nationalité		
N° de téléphone (facultatif)		
N° de télécopie (facultatif)		
Adresse électronique (facultatif)		
5 DEMANDEUR (Cochez l'une des 2 cases)		<input type="checkbox"/> Personne morale <input type="checkbox"/> Personne physique
Nom ou dénomination sociale		
Prénoms		
Forme juridique		
N° SIREN		
Code APE-NAF		
Domicile ou siège	Rue	
	Code postal et ville	
	Pays	
Nationalité		
N° de téléphone (facultatif)		
N° de télécopie (facultatif)		
Adresse électronique (facultatif)		
1 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE (Nom et qualité du signataire)		VISA DE LA PRÉFECTURE OU DE L'INPI
Eric BURBAUD 04 0304		

La loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire.
Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI

Procédé, dispositif informatique et programme d'ordinateur
pour l'aide à la voyellation de mots en langue arabe

L'invention concerne la voyellation d'un texte en langue
5 arabe, assistée par des moyens informatiques.

L'écriture de la langue arabe prévoit principalement deux
types de caractères. Un premier type concerne les
consonnes, lesquelles constituent le corps du texte. Un
10 second type concerne les voyelles, lesquelles, dans
l'écriture arabe, sont ajoutées aux consonnes par
adjonction de signes de voyellation au-dessus ou en-
dessous de chaque consonne.

15 Généralement, les textes publiés en langue arabe
comportent des mots représentés uniquement par leurs
consonnes. Seuls les ouvrages pédagogiques pour
l'apprentissage de la langue arabe présentent les
consonnes avec les signes de voyellation.

20

En se référant à la figure 1a, le mot représenté sur cette
figure comporte trois lettres successives 1, 2 et 3,
correspondant respectivement aux consonnes K, T et B. Ce
mot, dans son contexte, signifie habituellement "il a
25 écrit" et se lit KATABA. Un lecteur d'un texte arabe,
pratiquant couramment cette langue, interprétera donc
naturellement la succession des trois lettres de la figure
1a comme correspondant au mot KATABA, qui, lorsqu'il est
voyellé, présente des barres horizontales 4 figurant au-
30 dessus des lettres 1, 2 et 3, comme le montre la figure
1b. En se référant à la figure 1b, on comprendra ainsi que

ces barres horizontales 4, placées au-dessus des consonnes K, T, B, correspondent à la voyelle A et un lecteur non initié à la langue arabe peut déduire maintenant sans ambiguïté de l'expression représentée sur la figure 1b qu'il s'agit du mot KATABA.

Toutefois, en se référant à la figure 1c, le lecteur non initié ne saurait pas si le mot non voyellé de la figure la correspond :

- 10 - à la bonne combinaison de voyelles KATABA (portant la référence A sur la figure 1c),
- à la combinaison erronée de voyelles KATABO (portant la référence B sur la figure 1c),
- 15 - à la combinaison erronée de voyelles KOTOBO (portant la référence C sur la figure 1c),

ou à toute autre combinaison parmi 27 combinaisons possibles pour ces trois consonnes.

En effet, on peut compter en tout 9 signes possibles de voyellation pour une consonne (a, o, i, an, oun, in, pas de voyelle associée à la consonne, *hamza* et *chedda*).

Cette difficulté est d'autant accrue que certains mots, non voyellés, peuvent se lire selon une pluralité d'interprétations possibles. Par exemple, le mot "*homme*", non voyellé, se lit aussi bien "*homme*" que "*pied*", car le mot "*pied*", en arabe, présente la même succession de consonnes que le mot "*homme*".

Dans d'autres applications actuellement envisagées telles que la synthèse vocale (impliquant une conversion de caractères d'écriture en signaux de parole voisés), la

voyellation des mots apparaît nécessaire car une simple succession de consonnes ne peut pas permettre à elle seule la construction d'un signal de parole exact.

5 D'un autre côté, une voyellation manuelle d'un texte complet, édité électroniquement, est fastidieuse car l'opérateur doit actionner systématiquement une touche pour une consonne et au moins deux touches pour éditer en outre le signe de voyellation associée à cette consonne
10 (notamment la touche "SHIFT" et une autre touche du clavier).

Ainsi, il existe aujourd'hui un réel besoin d'une voyellation automatique de mots en langue arabe.

15

On connaît, à cet effet, un procédé assisté par des moyens informatiques et basé sur le découpage de mots en une pluralité de tronçons tels que, notamment, un préfixe, un radical, un suffixe. En suivant cet exemple, chaque type
20 de préfixe est stocké dans un premier dictionnaire, chaque type de radical est stocké dans un second dictionnaire et chaque type de suffixe est stocké dans un troisième dictionnaire. On procède de la même manière pour des verbes conjugués. Finalement, ce procédé prévoit une
25 multiplicité de dictionnaires formant des bases de données qui sont stockées dans une mémoire des moyens informatiques précités.

Ainsi, un mot à voyeller est découpé en plusieurs
30 tronçons. Chaque tronçon, comportant une succession identifiée de consonnes, est comparé avec une succession

de consonnes correspondante dans le dictionnaire qui est propre à ce type de tronçon. Des règles de voyellation codées sous la forme d'instructions de programme informatique définissent la voyellation qui doit être appliquée à ce tronçon. Finalement, le mot voyellé est reconstruit par concaténation des différents tronçons voyellés.

Ce procédé, quoique prometteur, présente de nombreuses erreurs dans sa mise en oeuvre. A titre illustratif, on comprendra par exemple que le mot "INFORMATION" comporte le radical "INFORM-" et le même suffixe "-ATION" que le mot "PERTURBATION". Toutefois, le mot "NATION" ne peut pas être découpé de la même manière avec la lettre seule "N-", d'une part, et la succession de lettres "-ATION", d'autre part. Le même problème se pose en langue arabe.

La présente invention vient améliorer la situation.

Se basant sur une toute autre approche, elle propose à cet effet un procédé de voyellation d'un texte en langue arabe, assisté par des moyens informatiques, dans lequel :

- a) on prévoit une première zone mémoire dans laquelle est stocké un premier dictionnaire comportant des mots non voyellés,
- b) on prévoit une seconde zone mémoire dans laquelle est stocké un second dictionnaire comportant des groupes d'au moins un mot voyellé, chaque groupe étant stocké en correspondance d'un mot non voyellé dudit premier dictionnaire,

c) pour un mot courant, non voyellé, on compare une chaîne de caractères formant au moins ledit mot courant avec des chaînes de caractères stockées dans la première zone mémoire, pour isoler au moins un mot du premier dictionnaire comportant la même chaîne de caractères que le mot courant, et

d) on extrait du second dictionnaire un groupe de mots candidats, voyellés, correspondant audit mot isolé du premier dictionnaire.

10

La présente invention vise aussi un dispositif informatique d'aide à la voyellation d'un texte en langue arabe, comportant :

- une première zone mémoire dans laquelle est stocké un premier dictionnaire comportant des mots non voyellés,

- une seconde zone mémoire dans laquelle est stocké un second dictionnaire comportant des groupes d'au moins un mot voyellé, chaque groupe étant stocké en correspondance d'un mot non voyellé dudit premier dictionnaire,

- une zone mémoire dans laquelle sont stockées des instructions d'une routine informatique propre à :

c) comparer, pour un mot courant, non voyellé, une chaîne de caractères formant au moins ledit mot courant avec des chaînes de caractères stockées dans la première zone mémoire, pour isoler au moins un mot du premier dictionnaire comportant la même chaîne de caractères que le mot courant, et

d) extraire du second dictionnaire un groupe de mots candidats, voyellés, correspondant audit mot isolé du premier dictionnaire.

30

A ce titre, la présente invention vise aussi un programme d'ordinateur d'aide à la voyellation d'un texte en langue arabe, stocké dans une mémoire d'un dispositif informatique ou, de manière équivalente, sur un support
5 destiné à coopérer avec un lecteur d'un dispositif informatique, comportant :

- une première base de données agencée selon un premier dictionnaire comportant des mots non voyellés,

- une seconde base de données agencée selon un second
10 dictionnaire comportant des groupes d'au moins un mot voyellé, chaque groupe de la seconde base étant indexé en correspondance d'un mot non voyellé de la première base, et

- une routine informatique propre à :

- 15 c) comparer, pour un mot courant, non voyellé, une chaîne de caractères formant au moins ledit mot courant avec des chaînes de caractères stockées dans la première zone mémoire, pour isoler au moins un mot du premier dictionnaire comportant la même chaîne de caractères que
20 le mot courant, et

- d) extraire du second dictionnaire un groupe de mots candidats, voyellés, correspondant audit mot isolé du premier dictionnaire.

25 On comprendra ainsi qu'une voyellation, au sens de l'invention, se base uniquement sur deux dictionnaires, l'un comprenant des mots non voyellés et l'autre, comprenant des groupes de mots voyellés. On verra dans la description d'un mode de réalisation préféré et de
30 variantes de ce mode de réalisation, donnée ci-après,

comment est sélectionné un mot candidat voyellé en remplacement d'un mot courant non voyellé.

5 D'autres caractéristiques et avantages de l'invention apparaîtront à l'examen de la description détaillée ci-après, et des dessins annexés sur lesquels :

- la figure 1a illustre un mot arabe non voyellé,
- la figure 1b illustre le mot de la figure 1a, mais voyellé maintenant,
- 10 - la figure 1c illustre le mot de la figure 1a, avec plusieurs voyellations possibles de ce mot,
- la figure 2 représente schématiquement un dispositif informatique pour la mise en œuvre de la présente invention,
- 15 - la figure 3 représente schématiquement le contenu de zones mémoires d'une mémoire de l'unité centrale 24 de la figure 2,
- les figures 4a, 4b et 4c représentent respectivement un texte comportant une phrase non voyellée, une phrase
- 20 voyellée sans voyelles casuelles et une phrase voyellée avec voyelles casuelles,
- la figure 5 représente un organigramme général du procédé selon un mode de réalisation préféré de l'invention,
- 25 - la figure 6 représente une boîte de dialogue mise en œuvre par un module d'interface homme/machine, pour proposer des voyellations possibles d'un mot courant, et
- la figure 7 représenté une boîte de dialogue proposant des étiquettes grammaticales possibles d'un mot courant.

On se réfère tout d'abord à la figure 2 sur laquelle un dispositif informatique comporte classiquement une unité centrale 24, à laquelle sont reliés un écran de visualisation 21, un organe de saisie tel qu'un clavier 22
5 ou une souris 23, ainsi qu'une interface de communication COM, par exemple avec un serveur distant, via un réseau étendu de type INTERNET. L'unité centrale 24 comporte en outre un lecteur 25 propre à coopérer avec un support mémoire tel qu'un CD-ROM, un DVD-ROM, une disquette, ou
10 tout autre support mémoire. On comprendra ainsi qu'un programme informatique, au sens de l'invention, peut être stocké sur un support mémoire de ce type, tandis que des mises à jour des dictionnaires précités peuvent être téléchargées du serveur distant ou encore obtenues sur un
15 autre support mémoire.

La figure 3 représente une structure d'une mémoire (par exemple de type ROM) dans laquelle sont stockés les premier et second dictionnaires précités. On indique que
20 l'unité centrale 24 comporte une mémoire, par exemple une mémoire permanente de type ROM, dans laquelle sont stockées sous forme numérique des successions de caractères arabes formant des mots des premier et second dictionnaires.

25

Une première zone mémoire D1 stocke un premier dictionnaire comportant des mots non voyellés 31, 32. Une seconde zone mémoire D2 stocke un second dictionnaire comportant des groupes 3-1, 3-2 d'un ou plusieurs mots
30 voyellés 311,312 ; 321,322. Préférentiellement, chaque groupe 3-1, 3-2 du second dictionnaire D2 est stocké en

correspondance d'un mot non voyellé 31, 32 du premier dictionnaire D1, comme l'illustrent les flèches de correspondance F11, F12, F21, F22 sur la figure 3. On retrouve par exemple dans le premier dictionnaire D1, la
5 succession des trois consonnes K,T,B (mot 31) de la figure 1a et, dans le second dictionnaire D2, le mot KATABA 311.

On indique que, dans une réalisation préférée, seuls les mots voyellés qui ont un sens sont répertoriés dans le
10 second dictionnaire précité. Toutefois, en variante, on peut prévoir de former un second dictionnaire initial comportant toutes les combinaisons possibles de voyelles pour une succession donnée de consonnes, tandis qu'un utilisateur supprime du second dictionnaire, au fur et à
15 mesure de l'utilisation, les combinaisons aberrantes et qui correspondent à des mots qui n'ont aucun sens. Dans ce cas, le second dictionnaire est formé par apprentissage en éliminant de la zone mémoire D2 les combinaisons aberrantes.

20

Cependant, dans le mode de réalisation préféré, le second dictionnaire est construit initialement avec des mots voyellés qui ont un sens, de manière à offrir une utilisation agréable et conviviale du programme au sens de
25 l'invention.

Bien entendu, pour un programme d'ordinateur d'aide à la voyellation au sens de l'invention, stocké dans une mémoire d'un dispositif informatique ou sur un support
30 capable de coopérer avec un lecteur d'un dispositif

informatique, les premier et second dictionnaires se présentent respectivement sous la forme :

- d'une première base de données D1 dont la structure est arrangée selon le premier dictionnaire qui comporte des mots non voyellés, et
- d'une seconde base de données D2 dont la structure est arrangée selon le second dictionnaire qui comporte des groupes d'au moins un mot voyellé.

Chaque groupe de la seconde base de données D2 est indexé en correspondance d'un mot non voyellé de la première base de données D1, comme le montrent encore les flèches de correspondance F11 à F22 de la figure 3.

On se réfère maintenant aux figures 4a et 4b qui représentent respectivement un texte non voyellé contenant une phrase complète délimitée par deux points P1 et P2 et un texte partiellement voyellé contenant ladite phrase délimitée par les points P1 et P2. On rappelle que l'arabe se lit de la droite vers la gauche. On comprendra ainsi qu'une succession de mots peut se présenter sous la forme d'une phrase complète définie par une chaîne de caractères entre deux caractères de ponctuation P1 et P2, les différents mots de cette phrase pouvant être voyellés en fonction de leur position dans la phrase, comme on le verra plus loin.

On indique simplement ici que le texte de la figure 4b ne comporte pas systématiquement de voyelles dites "casuelles" et qui sont attribuées le plus souvent en fin de mot. En revanche, le texte de la figure 4c est voyellé

de façon complète et comporte en outre les voyelles
casuelles qui apparaissent notamment à la dernière lettre
431 du mot 43 (avec un trait horizontal sous cette
dernière lettre 431 et à comparer avec la dernière lettre
5 non voyellée 421 du mot 42 (partiellement voyellé) de la
figure 4b).

En outre, on reconnaîtra sur la figure 4a le mot non
voyellé, référencé 45, qui comporte la succession de
10 caractère 1, 2, 3 de la figure 1a, correspondant aux
consonnes K, T, B. On reconnaîtra aussi sur la figure 4b
le mot voyellé 451 qui correspond au mot KATABA de la
figure 1b et voyellé par des traits horizontaux 4 au-
dessus des consonnes, qui sont représentatifs de la
15 voyelle "A".

Ces phrases des figures 4a, 4b et 4c apparaissent ainsi à
l'écran 21 du dispositif informatique et les caractères
des textes formant ces phrases sont classiquement stockés
20 sous forme numérique TXT (figure 3) dans une mémoire de
travail Z4 (par exemple de type RAM) de l'unité centrale
24 du dispositif informatique.

En se référant à nouveau à la figure 3, le dispositif
25 informatique comporte en outre une zone mémoire Z3 dans
laquelle sont stockées des instructions d'un programme
informatique PGM propre à :

- comparer, pour un mot courant non voyellé (portant la
référence 45 sur la figure 4a), une chaîne de caractères
30 (en l'espèce les consonnes 1, 2 et 3 de la figure 1a)
formant ce mot courant 45, avec des chaînes de caractères

31 stockées dans la première zone mémoire D1, pour isoler le mot 31 du premier dictionnaire D1 comportant la même chaîne de caractères que le mot courant 45, et
- extraire du second dictionnaire D2 un groupe 3-1 de mots
5 candidats 311, 321, voyellés et qui correspondent (flèches F11 et F12) au mot isolé 31 du premier dictionnaire D1.

On se réfère maintenant à la figure 5 pour décrire le déroulement de la routine informatique du programme PGM.
10 On cherche à voyeller ici un mot 45 qui apparaît dans un texte édité électroniquement sur l'écran 21 de la figure 2. Cette routine repère d'abord, par exemple par reconnaissance de caractères, à l'étape 51, les caractères (les consonnes 1, 2, 3) du mot non voyellé 45. La routine
15 effectue ensuite, à l'étape 52, une comparaison avec des mots non voyellés et répertoriés dans le dictionnaire D1 pour en isoler, à l'étape 53, un mot non voyellé 31 présentant la même succession de consonnes 1, 2, 3.

20 A l'étape 54, le programme PGM détermine, en fonction de l'emplacement mémoire dans la zone mémoire D1 du mot 31, l'emplacement mémoire du groupe 3-1 dans la zone mémoire D2 et comprenant les mots voyellés 311 et 312, du second dictionnaire de mots voyellés. A l'étape 55, le programme
25 PGM extrait de la zone mémoire D2 le groupe de mots candidats 311 et 312 comportant la même succession de consonnes mais voyellés différemment.

Dans un mode de réalisation préféré, on prévoit en outre
30 un module d'interface homme/machine, préférentiellement sous la forme d'instructions informatiques faisant partie

du programme PGM. Sur la figure 6, on montre une copie d'écran 21 faisant apparaître, pour un texte 62 édité électroniquement, une boîte de dialogue 61 qui est l'une des fonctionnalités de cette interface homme/machine. Pour un mot courant 45, non voyellé, sélectionné par un utilisateur (à partir d'un organe de saisie comme la souris 23) et qui apparaît, pour cette raison, contrasté dans le texte 62, la boîte de dialogue 61 indique d'abord quel est le mot 31 analysé en correspondance dans le premier dictionnaire D1. Ensuite, la boîte de dialogue 61 propose des voyellations potentielles de ce mot courant 45, qui correspondent à des mots voyellés candidats 312 et 311 du second dictionnaire D2, pour la même succession de consonnes que le mot 31 du premier dictionnaire. Ainsi, dans le deuxième cadre de la boîte de dialogue 61, l'interface homme/machine propose à un utilisateur une liste de choix des mots candidats 311 et 312.

En se référant à nouveau à la figure 5, dans un mode de réalisation préférée, l'utilisateur choisit, à l'étape 56, un mot candidat 311 parmi la liste de mots candidats 311, 312 du groupe de mots 3-1. A l'étape 57, le mot choisi 311, voyellé, remplace automatiquement le mot 45, non voyellé, dans le texte édité électroniquement. On précise en outre que "le choix" de l'utilisateur est mémorisé à l'étape 58, dans une zone mémoire Z5 du dispositif informatique. Préférentiellement, cette zone mémoire Z5 est en correspondance de la zone mémoire D2 dans laquelle est stocké le second dictionnaire, de manière à enrichir ce dernier. Plus particulièrement, le mot choisi 311, ainsi voyellé, est stocké avec les mots le précédant et/ou

le succédant dans une partie du texte édité. Préférentiellement, le mot choisi 311 est stocké avec la phrase complète dans laquelle il apparaît, en vue de perfectionner la voyellation au sens de la présente invention, par apprentissage, comme on le verra plus loin. On indique simplement ici que, si le mot courant 45 à voyeller fait partie d'une succession courante de mots, telle qu'une phrase complète, suite au choix d'un mot 311 par l'utilisateur (dans la liste de mots candidats 311, 312), le mot voyellé choisi 311 et la succession de mots qui le comporte sont mémorisés dans la zone mémoire Z5 précitée.

Ainsi, dans le troisième cadre de la boîte de dialogue 61 de la figure 6, l'interface homme/machine indique à l'utilisateur le mot choisi 311, qui va être édité dans le texte 62 en remplacement du mot 45 non voyellé et préférentiellement mémorisé avec une succession de mots le précédant et/ou le succédant.

On se réfère à nouveau aux figures 4a à 4c pour décrire ci-après une voyellation des mots en fonction de leur contexte.

Sur la figure 4a, on s'intéresse en particulier au premier mot de la phrase qui suit le point P1, sachant que l'arabe se lit de la droite vers la gauche. On reconnaît ce premier mot de la phrase sur la figure 3 qui correspond à l'expression non voyellée 32 du premier dictionnaire D1. Or, ce mot non voyellé 32 admet deux voyellations possibles 321 (signifiant l'expression "il est allé") et

322 (signifiant le métal "or") dans le second dictionnaire D2.

5 Généralement, en langue arabe, un mot commençant une phrase correspond à un verbe. Ainsi, le mot qui succède au premier point P1 de la figure 4a est un verbe dont la forme voyellée correspond avec quasi-certitude au verbe conjugué 321 du second dictionnaire D2 de la figure 3.

10 Ainsi, si le mot courant fait partie d'une succession de mots, on compare, de façon plus large, une chaîne de caractères formant cette succession de mots comportant le mot courant, avec des chaînes de caractères stockées dans la zone Z5 précitée en correspondance de la seconde zone
15 mémoire D2, pour identifier une pluralité de mots comportant une même chaîne de caractères que cette succession de mots. Cette étape correspond, dans une perspective plus large, à l'étape 51 représentée sur la figure 5.

20 On indique alors que le programme PGM peut comporter des instructions pour effectuer cette comparaison "élargie à une succession de mots". Par exemple, pour une phrase complète, une routine informatique peut être prévue pour
25 isoler les caractères de la phrase complète entre les deux signes de ponctuation P1 et P2.

Ensuite, pour le mot courant à voyeller, on sélectionne parmi le groupe de mots candidats voyellés extraits du
30 second dictionnaire D2, un mot voyellé (ici le verbe 321) en fonction de la succession de mots identifiés et, en

particulier, d'une position du mot courant 32 dans cette succession de mots identifiés. Ici, le mot 32 commence la phrase et correspond donc au verbe voyellé 321.

5 Avantageusement, on peut procéder alors à un remplacement automatique, dans le texte édité électroniquement, du mot courant non voyellé 32 par le mot voyellé 321, sélectionné automatiquement parmi le groupe de mots candidats 321 et 322.

10

On comprendra ainsi que cette voyellation automatique est avantageusement assurée ici par mémorisation de phrases complètes et/ou de succession de mots, dont la voyellation est validée par l'utilisateur, au fur et à mesure de
15 l'utilisation du logiciel informatique d'aide à la voyellation, donc par apprentissage. Des routines d'apprentissage informatique sont connues en soi. On indique par exemple que des routines telles que celles utilisées par le logiciel ViaVoice ® de la société
20 Microsoft ® conviennent bien à la détermination de caractères écrits par apprentissage.

Toutefois, en cas d'incertitude sur la voyellation, l'interface homme/machine propose avantageusement à
25 l'utilisateur une liste de choix comportant des mots sélectionnés parmi des mots candidats du second dictionnaire. Cette situation est représentée sur la figure 6 où deux voyellations possibles 312 et 311, qui sont cohérentes en fonction du contexte du mot courant 45,
30 sont proposées à l'utilisateur. De façon encore plus avantageuse, cette liste est hiérarchisée, en fonction du

contexte, par ordre de pertinence des voyellations proposées. En particulier, cette hiérarchie peut être déduite par apprentissage, en analysant la forme de voyellation préférée par l'utilisateur et qui revient le plus souvent au cours de l'utilisation.

En se référant à la figure 7, de façon avantageuse, des étiquettes grammaticales en correspondance de chaque mot 311 dans chaque groupe 3-1 du second dictionnaire D2 sont stockées dans une zone mémoire (non représentée), de sorte que l'interface homme/machine, en particulier la boîte de dialogue 61 de la figure 7, indique en outre à l'utilisateur une étiquette grammaticale 70 de chacun des mots sélectionnés parmi les mots candidats 311, 312. Le cas échéant, cette étiquette grammaticale est validée par l'utilisateur, dans le cadre 71 de la boîte de dialogue. On indique que cette étiquette grammaticale correspond par exemple à une description syntaxique d'un mot, du type "*nom commun, au singulier, défini, placé en tant que sujet dans la phrase, etc*". Bien entendu, cette étiquette grammaticale est définie et validée en fonction de la position du mot analysé 45 dans la phrase courante.

A cet effet, on prévoit une zone mémoire (par exemple en correspondance encore de la seconde zone mémoire D2) pour stocker en outre des étiquettes grammaticales 70 correspondant chacune à un mot voyellé 311 du second dictionnaire.

Comme le montrent les figures 6 et 7, on précise que le programme informatique PGM, pour la mise en œuvre de

l'invention, ainsi que le module d'interface homme/machine, sont compatibles avec des moyens électroniques d'édition de texte en langue arabe, tels que le logiciel *MICROSOFT WORD*®.

5

On décrit ci-après un autre type de voyellation automatique possible, dite "casuelle". Les voyelles casuelles sont le plus souvent attribuées aux consonnes de fin de mot, suivant le contexte de ce mot dans une phrase.

10 Par exemple, le mot 42 de la figure 4b, dans son contexte, admet une voyellation de sa dernière lettre 421, par le son "i" qui correspond à une barre horizontale 431 sous cette lettre de fin.

15 On rappelle qu'il existe, dans la langue arabe, une pluralité de déclinaisons possibles pour un nom commun, telles que le nominatif (déterminé ou indéterminé), l'accusatif (déterminé ou indéterminé), l'ablatif (déterminé ou indéterminé), etc. A ces déclinaisons
20 correspondent des voyellations de fin de mot avec les sons suivants :

- "O" = nominatif déterminé,
- "OUN" = nominatif indéterminé,
- "A" = accusatif déterminé,
- 25 - "AN" = accusatif indéterminé,
- "I" = ablatif déterminé,
- "IN" = ablatif indéterminé, etc.

Par exemple, en se référant à nouveau aux figures 4b et
30 4c, la préposition correspondant au mot 44 est repérée dans la succession de mots dans laquelle figure le mot 43.

Cette préposition 44 entraîne nécessairement une déclinaison à l'ablatif du mot 43 qui suit, avec une voyellation casuelle automatique par le son "i" de la dernière lettre 431 du mot 43.

5

Ainsi, comme précédemment, la routine informatique du programme PGM comporte des instructions pour comparer la succession courante de mots de la figure 4b, avec des successions de mots mémorisées au préalable. Le cas échéant, la préposition 44 est identifiée, avec une position qui précède juste le mot 42 à voyeller. Une routine du programme PGM sélectionne alors, en fonction de cette comparaison, le mot voyellé 43 finissant par le son "i" qui correspond à une déclinaison à l'ablatif, entraînée par la position de cette préposition 44 par rapport au mot 43. On indique que la voyellation casuelle est proposée comme option par l'interface homme/machine du programme PGM, dans un mode de réalisation préféré.

De manière générale, on comprendra que les étapes décrites ci-avant, notamment celles en référence avec la figure 5, sont mise en œuvre par le déroulement d'instructions ou de routines informatiques du programme PGM, lequel est donc destiné à être installé dans une mémoire d'une machine ou d'un dispositif informatique du type représenté sur la figure 2. Initialement, ce programme, par exemple stocké sur CD-ROM, comporte les première et seconde zones mémoires D1 et D2 agencées sous forme de bases de données (avec, le cas échéant, les données des étiquettes grammaticales), lesquelles peuvent être chargées et copiées en mémoire (par exemple permanente de type ROM) du

30

dispositif informatique précité. On comprendra que ces bases de données, une fois copiées en mémoire du dispositif, peuvent ensuite être enrichies, notamment par apprentissage. En particulier, il en va de même pour
5 ladite zone mémoire Z5 en correspondance de la seconde zone mémoire, qui est destinée à stocker les successions de mots ou des phrases complètes. La base de données stockée dans la zone Z5 (dans une mémoire du dispositif) est ainsi enrichie au fur et à mesure de l'utilisation.

Revendications

1. Procédé de voyellation d'un texte en langue arabe, assisté par des moyens informatiques, dans lequel :
- 5 a) on prévoit une première zone mémoire dans laquelle est stocké un premier dictionnaire comportant des mots non voyellés,
- b) on prévoit une seconde zone mémoire dans laquelle est stocké un second dictionnaire comportant des groupes d'au
- 10 moins un mot voyellé, chaque groupe étant stocké en correspondance d'un mot non voyellé dudit premier dictionnaire,
- c) pour un mot courant, non voyellé, on compare une chaîne de caractères formant au moins ledit mot courant avec des
- 15 chaînes de caractères stockées dans la première zone mémoire, pour isoler au moins un mot du premier dictionnaire comportant la même chaîne de caractères que le mot courant, et
- d) on extrait du second dictionnaire un groupe de mots
- 20 candidats, voyellés, correspondant audit mot isolé du premier dictionnaire.
2. Procédé selon la revendication 1, dans lequel on prévoit une routine informatique propre à effectuer ladite
- 25 comparaison des chaînes de caractères et ladite extraction du groupe de mots candidats.
3. Procédé selon la revendication 1, dans lequel on prévoit en outre une interface homme/machine propre à
- 30 proposer à un utilisateur une liste de choix desdits mots candidats.

4. Procédé selon la revendication 1, dans lequel, ledit mot courant faisant partie d'une succession de mots,
- 5 c1) on compare une chaîne de caractères formant ladite succession de mots comportant le mot courant, avec des chaînes de caractères stockées dans une zone mémoire en correspondance de la seconde zone mémoire, pour identifier une pluralité de mots comportant une même chaîne de caractères que ladite succession de mots, et
- 10 d2) pour ledit mot courant, on sélectionne parmi ledit groupe de mots candidats voyellés, au moins un mot voyellé en fonction de la succession de mots identifiés et d'une position du mot courant dans ladite succession de mots identifiés.
- 15
5. Procédé selon la revendication 4, dans lequel ladite succession de mots est une phrase complète définie par une chaîne de caractères entre deux caractères de ponctuation.
- 20 6. Procédé selon la revendication 4, dans lequel on remplace automatiquement dans un texte édité électroniquement ledit mot courant par ledit mot voyellé, sélectionné parmi le groupe de mots candidats.
- 25 7. Procédé selon la revendication 3 et la revendication 4, dans lequel l'interface homme/machine propose à un utilisateur une liste de choix comportant des mots sélectionnés parmi lesdits mots candidats.
- 30 8. Procédé selon la revendication 7, dans lequel on stocke en outre des étiquettes grammaticales en correspondance de

chaque mot dans chaque groupe du second dictionnaire, et dans lequel l'interface homme/machine indique en outre à l'utilisateur une étiquette grammaticale de chacun des mots sélectionnés parmi lesdits mots candidats.

5

9. Procédé selon la revendication 3, dans lequel, ledit mot courant faisant partie d'une succession courante de mots,

10 suite au choix d'un mot par ledit utilisateur dans la liste de mots candidats, on mémorise le mot choisi avec ladite succession de mots, dans une zone mémoire en correspondance de ladite seconde zone mémoire.

10. Procédé selon la revendication 8 et la revendication 15 4, dans lequel la sélection du mot voyellé parmi ledit groupe de mots candidats voyellés est effectuée par apprentissage, en comparant la succession courante de mots avec des successions de mots mémorisées dans ladite zone mémoire en correspondance de la seconde zone mémoire.

20

11. Dispositif informatique d'aide à la voyellation d'un texte en langue arabe, comportant :

- une première zone mémoire dans laquelle est stocké un premier dictionnaire comportant des mots non voyellés,
- 25 - une seconde zone mémoire dans laquelle est stocké un second dictionnaire comportant des groupes d'au moins un mot voyellé, chaque groupe étant stocké en correspondance d'un mot non voyellé dudit premier dictionnaire,
- une zone mémoire dans laquelle sont stockées des
- 30 instructions d'une routine informatique propre à :

c) comparer, pour un mot courant, non voyellé, une chaîne de caractères formant au moins ledit mot courant avec des chaînes de caractères stockées dans la première zone mémoire, pour isoler au moins un mot du premier dictionnaire comportant la même chaîne de caractères que le mot courant, et

d) extraire du second dictionnaire un groupe de mots candidats, voyellés, correspondant audit mot isolé du premier dictionnaire.

10

12. Dispositif informatique selon la revendication 11, comportant en outre une interface homme/machine propre à proposer à un utilisateur une liste de choix desdits mots candidats.

15

13. Dispositif informatique selon la revendication 11, dans lequel, ledit mot courant faisant partie d'une succession de mots, ladite routine informatique est agencée pour :

20

c1) comparer une chaîne de caractères formant ladite succession de mots comportant le mot courant, avec des chaînes de caractères stockées dans une zone mémoire en correspondance de la seconde zone mémoire, pour identifier une pluralité de mots comportant une même chaîne de caractères que ladite succession de mots, et

25

d2) pour ledit mot courant, sélectionner parmi ledit groupe de mots candidats voyellés, au moins un mot voyellé en fonction de la succession de mots identifiés et d'une position du mot courant dans ladite succession de mots identifiés.

30

14. Dispositif informatique selon la revendication 13, dans lequel ladite succession de mots est une phrase complète définie par une chaîne de caractères entre deux caractères de ponctuation, et dans lequel ladite routine
5 informatique est agencée pour isoler les caractères de la phrase complète entre les deux signes de ponctuation.

15. Dispositif informatique selon la revendication 11, comportant en outre des moyens électroniques d'édition de
10 texte en langue arabe, et dans lequel ladite routine informatique est apte à coopérer avec lesdits moyens d'édition de texte.

16. Dispositif informatique selon la revendication 15 et
15 la revendication 13, dans lequel la routine informatique est agencée pour remplacer automatiquement dans un texte édité ledit mot courant par ledit mot voyellé, sélectionné parmi le groupe de mots candidats.

20 17. Dispositif informatique selon la revendication 12 et la revendication 13, dans lequel l'interface homme/machine est agencée pour proposer une liste de choix comportant des mots sélectionnés parmi lesdits mots candidats.

25 18. Dispositif informatique selon la revendication 12, dans lequel, ledit mot courant faisant partie d'une succession courante de mots,
la routine informatique comporte en outre des instructions pour mémoriser le mot choisi avec ladite succession de
30 mots, dans une zone mémoire en correspondance de ladite seconde zone mémoire.

19. Dispositif informatique selon la revendication 18 et la revendication 13, dans lequel la routine informatique comporte des instructions pour comparer la succession
5 courante de mots avec des successions de mots mémorisées dans ladite zone mémoire en correspondance de la seconde zone mémoire, et sélectionner, en fonction de cette comparaison, au moins un mot voyellé parmi ledit groupe de mots candidats voyellés.

10

20. Dispositif informatique selon la revendication 17, comportant une zone mémoire pour stocker en outre des étiquettes grammaticales en correspondance de chaque mot dans chaque groupe du second dictionnaire, et dans lequel
15 l'interface homme/machine indique en outre à l'utilisateur une étiquette grammaticale de chacun des mots sélectionnés parmi lesdits mots candidats.

21. Programme d'ordinateur d'aide à la voyellation d'un
20 texte en langue arabe, stocké dans une mémoire d'un dispositif informatique ou sur un support destiné à coopérer avec un lecteur d'un dispositif informatique, comportant :

- une première base de données agencée selon un premier
25 dictionnaire comportant des mots non voyellés,
- une seconde base de données agencée selon un second dictionnaire comportant des groupes d'au moins un mot voyellé, chaque groupe de la seconde base étant indexé en correspondance d'un mot non voyellé de la première base,
30 et
- une routine informatique propre à :

c) comparer, pour un mot courant, non voyellé, une chaîne de caractères formant au moins ledit mot courant avec des chaînes de caractères stockées dans la première zone mémoire, pour isoler au moins un mot du premier dictionnaire comportant la même chaîne de caractères que le mot courant, et

d) extraire du second dictionnaire un groupe de mots candidats, voyellés, correspondant audit mot isolé du premier dictionnaire.

10

22. Programme d'ordinateur selon la revendication 21, destiné à être installé dans une mémoire d'une machine informatique et comportant un module d'interface homme/machine propre à proposer à un utilisateur une liste de choix desdits mots candidats.

15

23. Programme d'ordinateur selon la revendication 21, dans lequel, ledit mot courant faisant partie d'une succession de mots, le programme comprend des instructions pour :

20

c1) comparer une chaîne de caractères formant ladite succession de mots comportant le mot courant, avec des chaînes de caractères stockées dans une zone mémoire en correspondance de la seconde zone mémoire, pour identifier une pluralité de mots comportant une même chaîne de caractères que ladite succession de mots, et

25

d2) pour ledit mot courant, sélectionner parmi ledit groupe de mots candidats voyellés, au moins un mot voyellé en fonction de la succession de mots identifiés et d'une position du mot courant dans ladite succession de mots identifiés.

30

24. Programme d'ordinateur selon la revendication 23, dans lequel ladite succession de mots est une phrase complète définie par une chaîne de caractères entre deux caractères de ponctuation, et dans lequel le programme comporte des instructions pour isoler les caractères de la phrase
5 complète entre les deux signes de ponctuation.

25. Programme d'ordinateur selon la revendication 21, compatible et apte à coopérer avec un programme d'édition
10 de texte en langue arabe.

26. Programme d'ordinateur selon la revendication 25 et la revendication 23, destiné à être installé dans une mémoire d'un dispositif informatique et comportant des
15 instructions pour remplacer automatiquement dans un texte édité ledit mot courant par ledit mot voyellé, sélectionné parmi le groupe de mots candidats.

27. Programme d'ordinateur selon la revendication 22 et la
20 revendication 23, dans lequel le module d'interface homme/machine est agencé pour proposer une liste de choix comportant des mots sélectionnés parmi lesdits mots candidats.

28. Programme d'ordinateur selon la revendication 22, dans lequel, ledit mot courant faisant partie d'une succession courante de mots,
le programme d'ordinateur comporte en outre des instructions pour mémoriser le mot choisi avec ladite
25 succession de mots, dans une zone mémoire en correspondance de ladite seconde zone mémoire.
30

29. Programme d'ordinateur selon la revendication 28 et la revendication 23, dans lequel le programme informatique comporte des instructions pour comparer la succession courante de mots avec des successions de mots mémorisées dans ladite zone mémoire en correspondance de la seconde zone mémoire, et sélectionner, en fonction de cette comparaison, au moins un mot voyellé parmi ledit groupe de mots candidats voyellés.

10

30. Programme d'ordinateur selon la revendication 27, comportant une base de données stockées en correspondance de chaque mot du second dictionnaire et comportant des étiquettes grammaticales pour chaque mot dans chaque groupe du second dictionnaire, dans lequel l'interface homme/machine comporte des instructions pour indiquer en outre à l'utilisateur une étiquette grammaticale de chacun des mots sélectionnés parmi lesdits mots candidats.

20

1/6

FIG. 1a

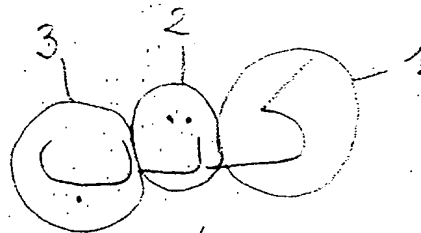


FIG. 1b

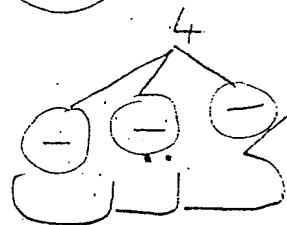


FIG. 1c

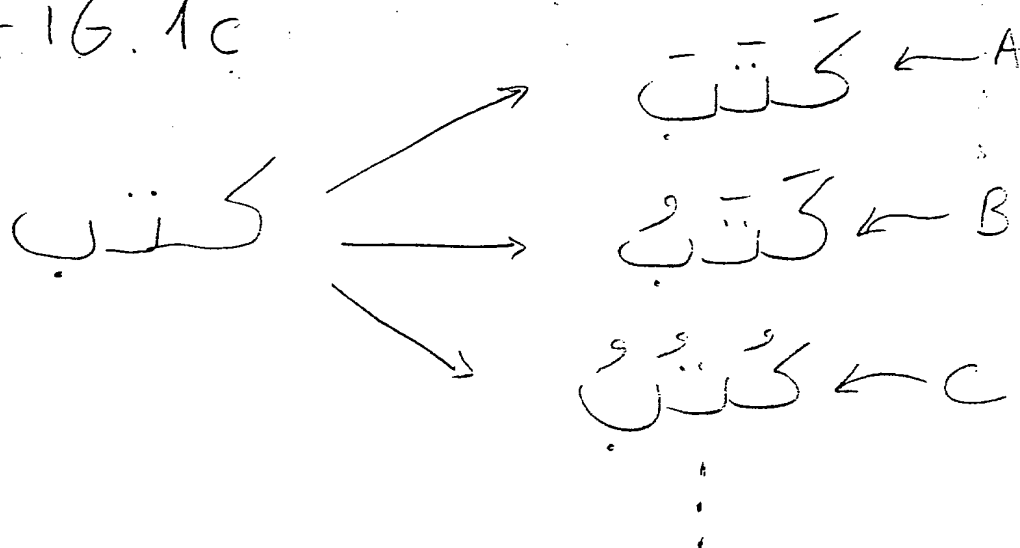


FIG. 1a

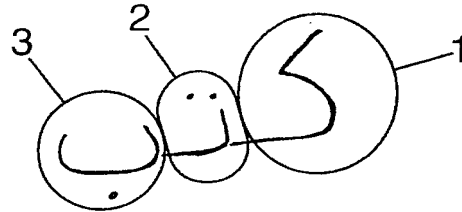


FIG. 1b

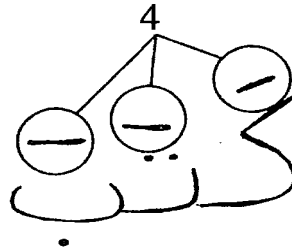
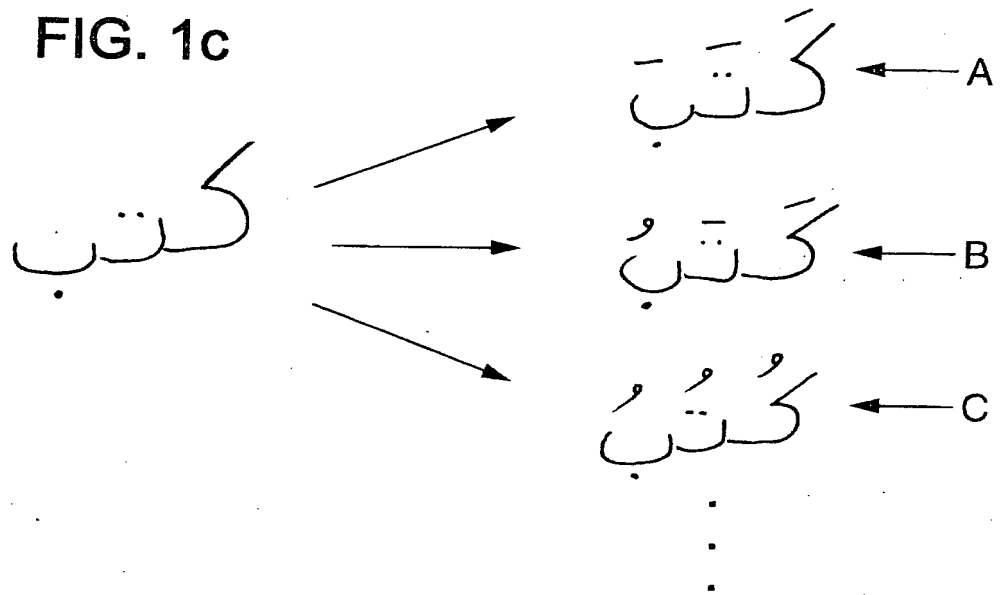


FIG. 1c



2/6

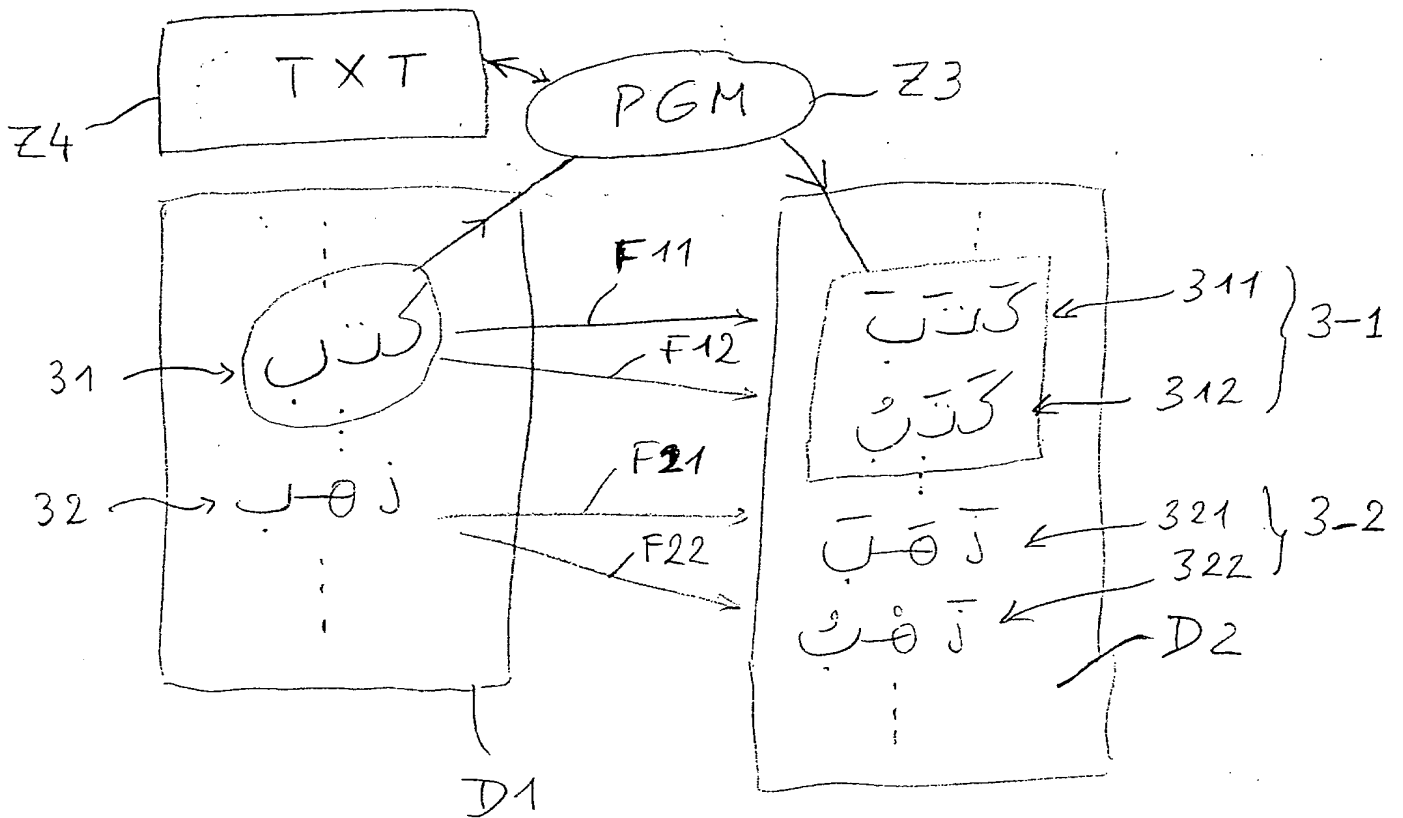
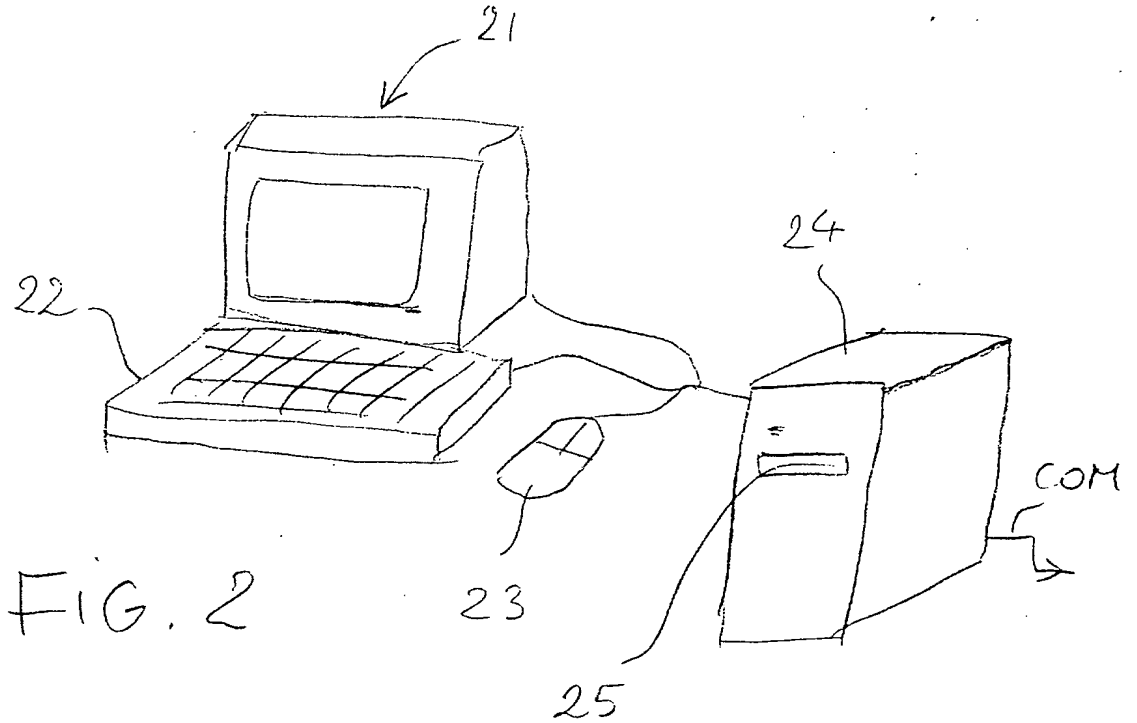


FIG. 2

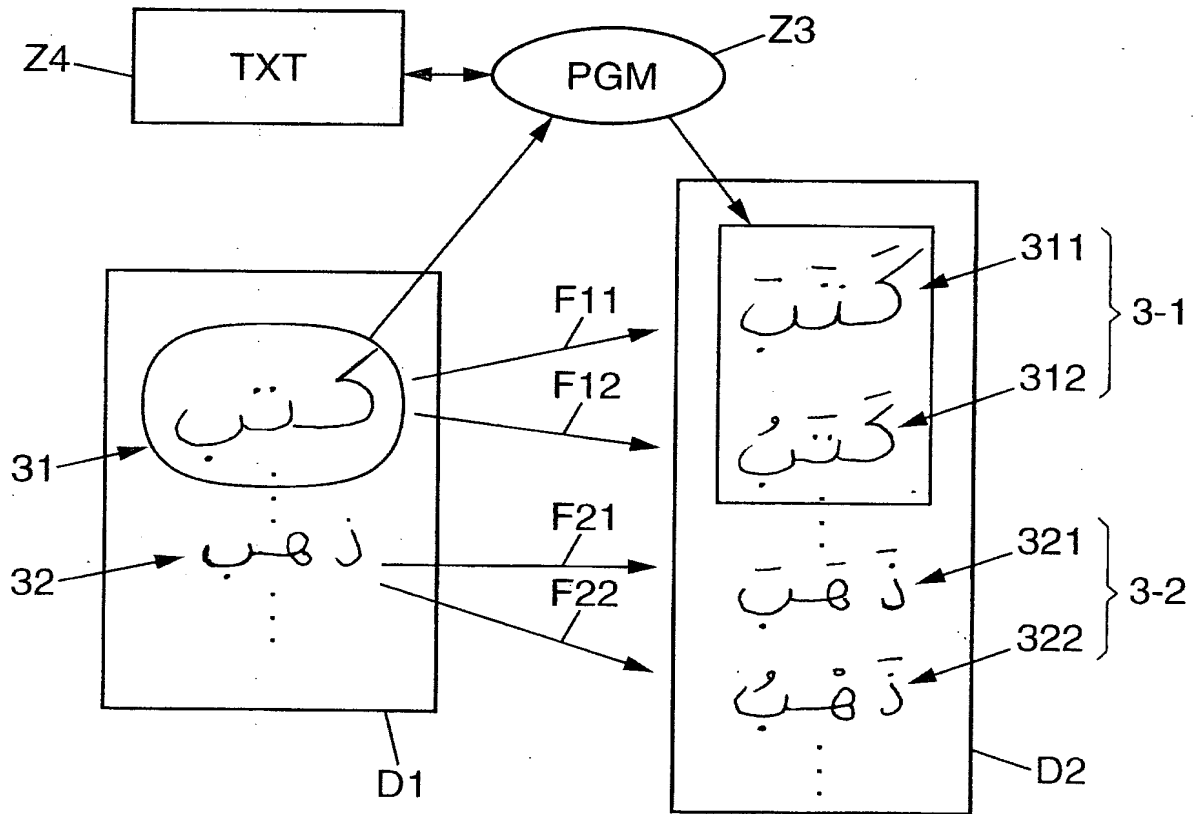
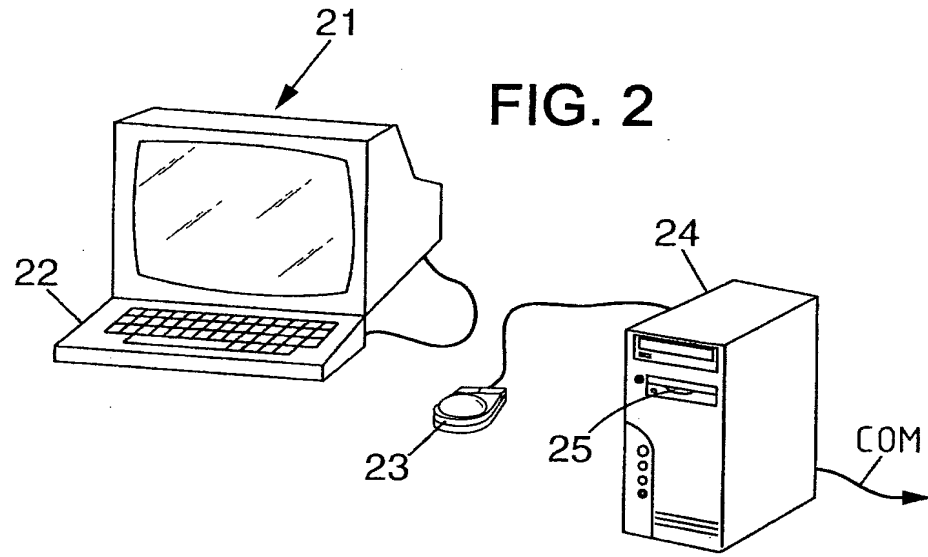


FIG. 3

...
 المدبنة الذهب محمد الى P1
 P2 بالقللا م ففقال ...
 المدرسة 41
 45

FIG. 4a

...
 المدبنة الذهب محمد الى P1
 P2 بالقللا م ففقال ...
 المدرسة 42
 451 421

FIG. 4b

...
 المدبنة الذهب محمد الى 44
 بالقللا م ففقال ...
 المدرسة 43
 431

FIG. 4c

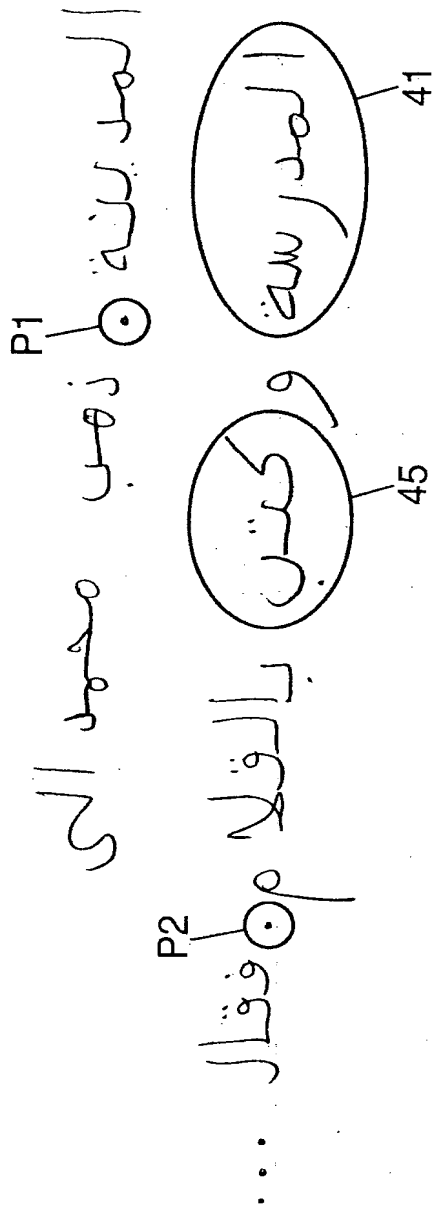


FIG. 4a

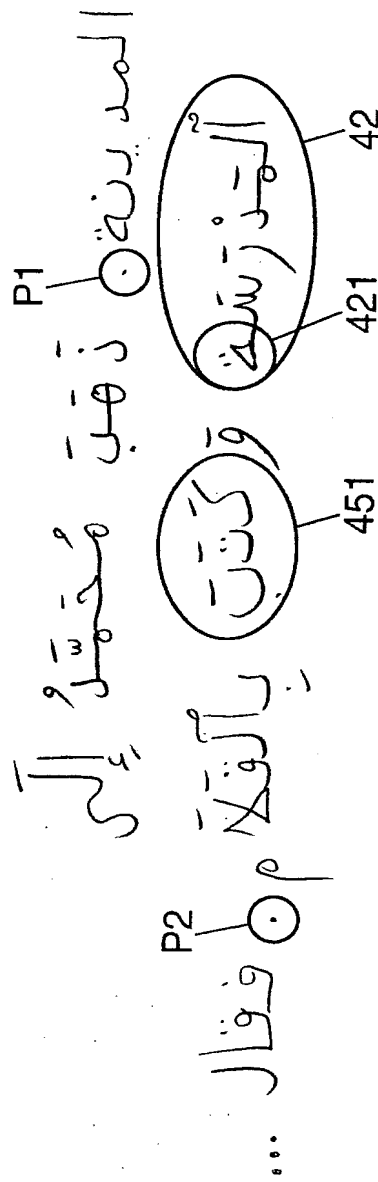


FIG. 4b

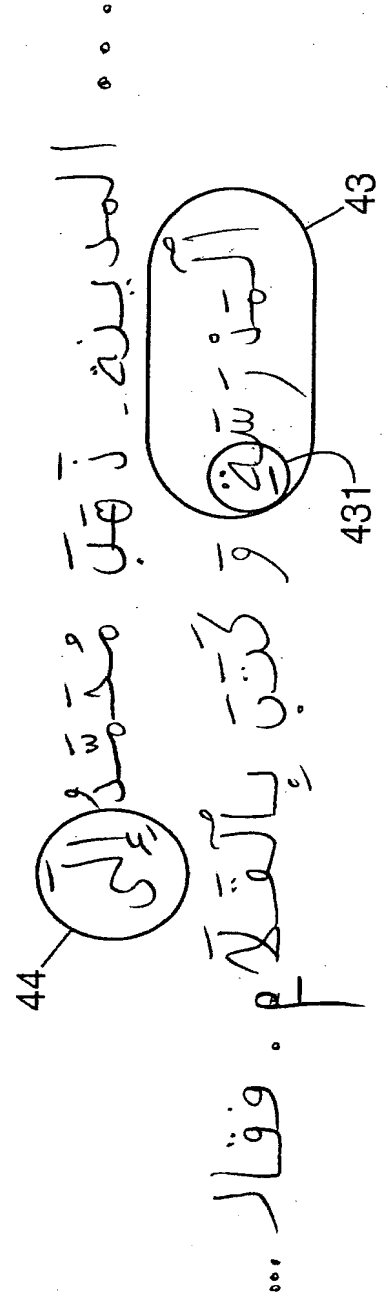
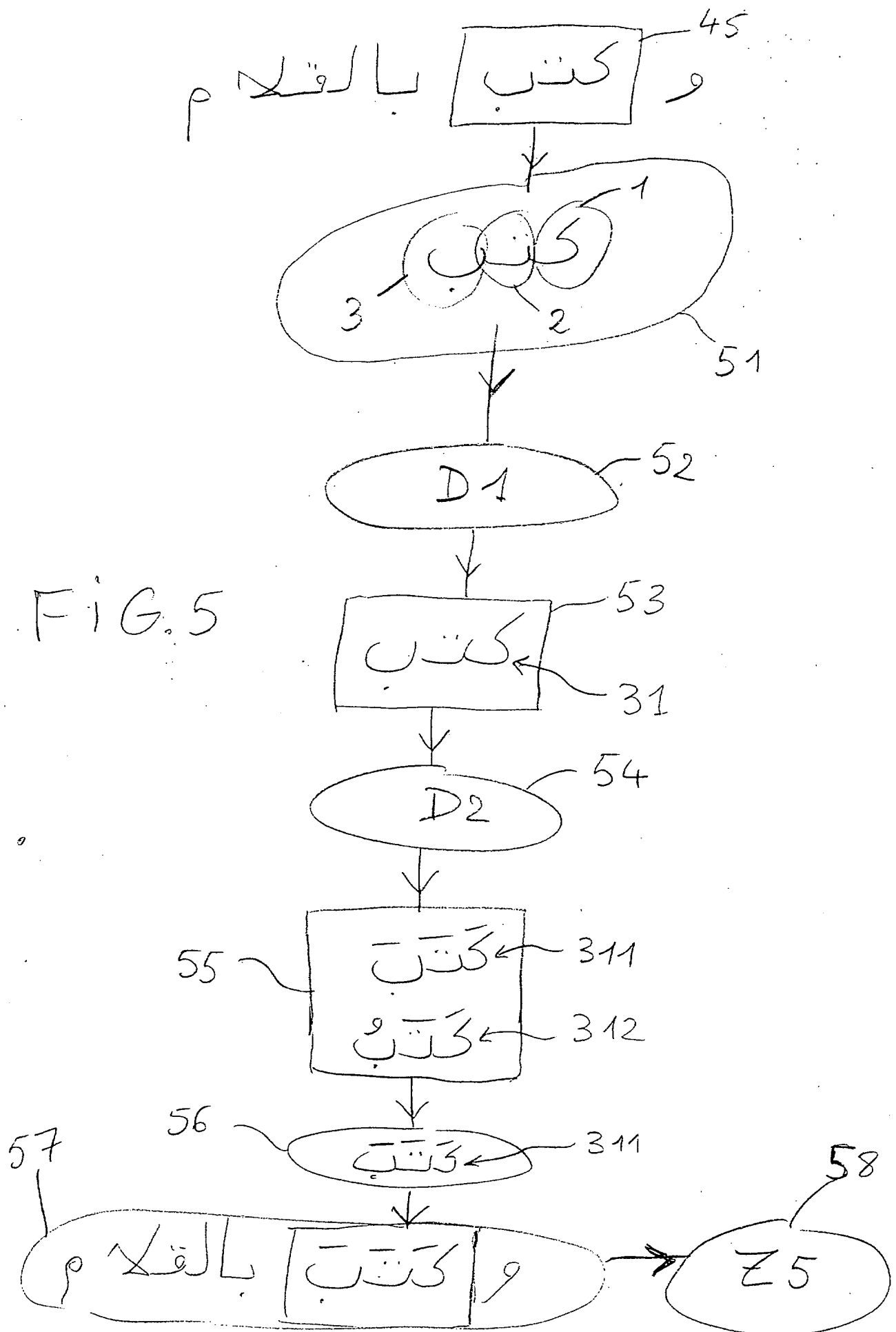
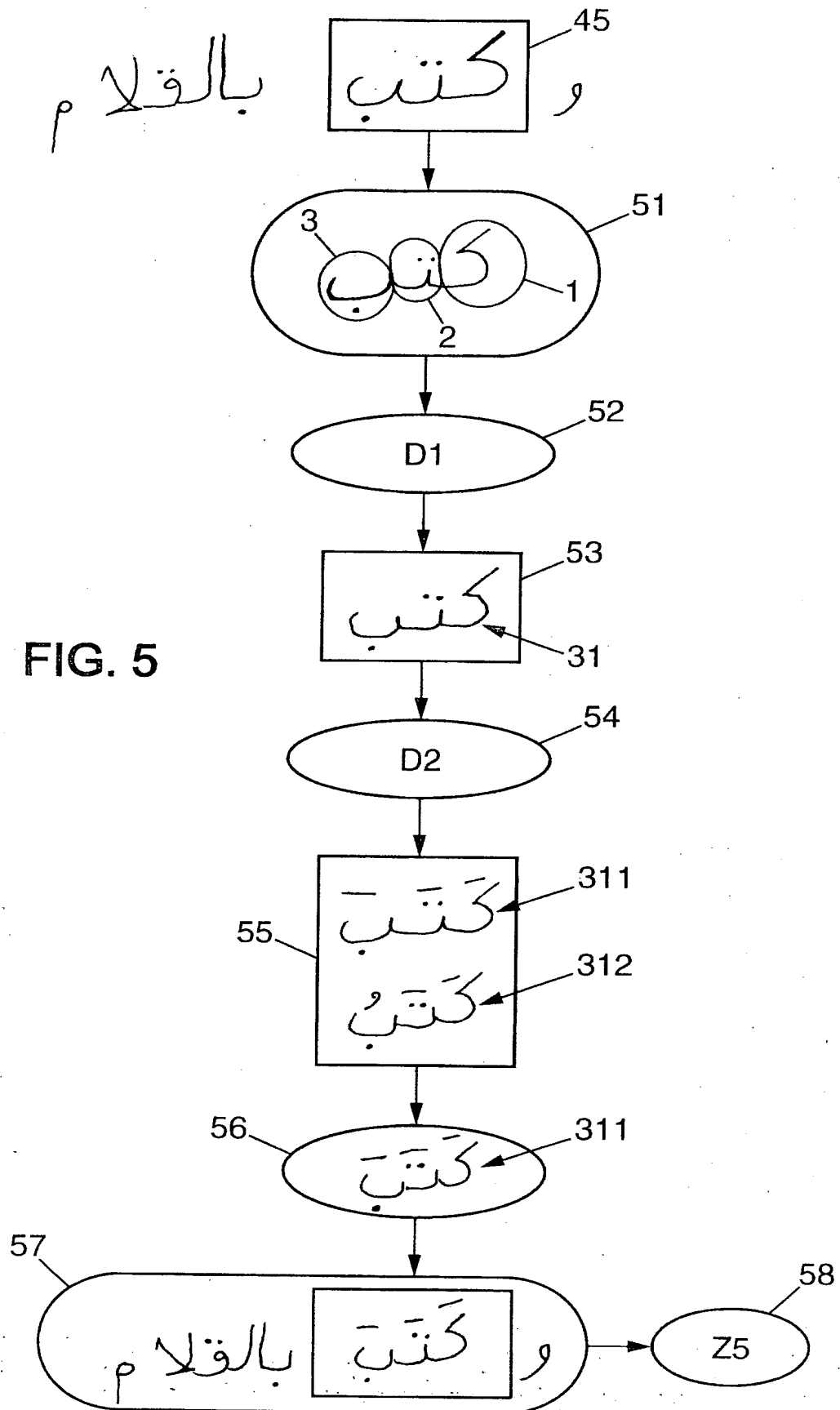


FIG. 4c



4/6





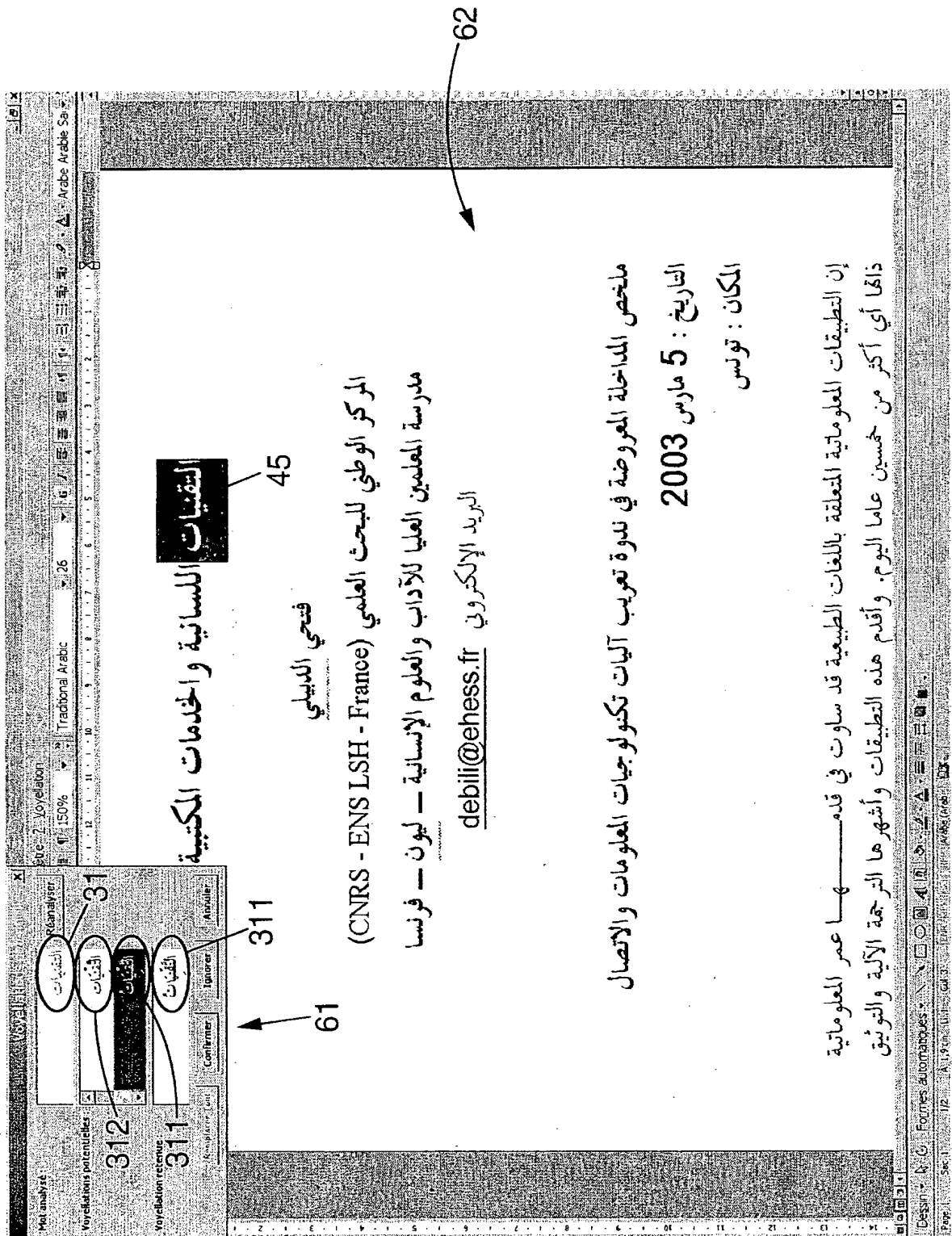


FIG. 6

6/6

62

Etiquetage automatique

Mot analysé: 311

Voyelles retenues: 31

Catégories potentielles: 70

Catégorie retenue: 71

Amidat

Ignorer

Confirmer

Remplacer Text

الفتيات اللسانية والفتيات

البريد الإلكتروني debili@ehess.fr

المركز الوطني للبحث العلمي (CNRS - ENS LSH - France)

مدرسة المعلمين العليا للأدب والعلوم الإنسانية — ليون — فرنسا

فتحي الديلي

ملخص المداخلة المروضة في ندوة تقريب آليات تكنولوجيات المعلومات والاتصال

التاريخ : 5 مارس 2003

المكان : تونس

إن التطبيقات المعلوماتية المتعلقة باللغات الطبيعية قد ساوت في قدمها عصر المعلوماتية ذاتها أي أكثر من خمسين عاما اليوم. وأقدم هذه التطبيقات وأشهرها الترجمة الآلية والتوثيق

Page 1 Sec 1 1/2 A 3.4 cm U 2 Col 1

Formes automatiques

Page 1

Fig. 7

Etiquette grammaticale

Arabe

25

31

التقنيات

311

التقنيات

70

إسم معرف ب "أ"

71

إسم معرف ب "أ"

61

فتحي الديبيلي

62

التقنيات اللسانية و

45

المركز الوطني للبحث العلمي (CNRS - ENS LSH - France)

مدرسة المعلمين العليا للأدب والعلوم الإنسانية — ليون — فرنسا

البريد الإلكتروني debili@ehess.fr

ملخص المداخلة المعروضة في ندوة تعريب آليات تكنولوجيا المعلومات والاتصال

التاريخ : 5 مارس 2003

المكان : تونس

إن التطبيقات المعلوماتية المتعلقة باللغات الطبيعية قد ساوت في قدمها عمر المعلوماتية ذاتها أي أكثر من خمسين عاما اليوم. وأقدم هذه التطبيقات وأشهرها الترجمة الآلية والتوثيق

Formes automatiques

Page 1

FIG. 7

**BREVET D'INVENTION****CERTIFICAT D'UTILITÉ**

Code de la propriété intellectuelle - Livre VI



N° 11235*03

DÉPARTEMENT DES BREVETS

26 bis, rue de Saint Pétersbourg

75800 Paris Cedex 08

Téléphone : 33 (1) 53 04 53 04 Télécopie : 33 (1) 42 94 86 54

DÉSIGNATION D'INVENTEUR(S) Page N° 1.1.1.

(À fournir dans le cas où les demandeurs et les inventeurs ne sont pas les mêmes personnes)



Cet imprimé est à remplir lisiblement à l'encre noire

DB 113 @ W / 270601

Vos références pour ce dossier (facultatif)		BFF030245	
N° D'ENREGISTREMENT NATIONAL		0307665	
TITRE DE L'INVENTION (200 caractères ou espaces maximum)			
Procédé, dispositif informatique et programme d'ordinateur pour l'aide à la voyellation de mots en langue arabe.			
LE(S) DEMANDEUR(S) :			
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE - CNRS - ECOLE NORMALE SUPERIEURE LETTRE ET SCIENCES HUMAINES			
DESIGNE(NT) EN TANT QU'INVENTEUR(S) :			
1 Nom			
Prénoms		Fathi DEBILI	
Adresse	Rue	21, rue Boris Vildé 92260 FONTENAY AUX ROSES FRANCE	
	Code postal et ville		
Société d'appartenance (facultatif)			
2 Nom			
Prénoms			
Adresse	Rue		
	Code postal et ville		
Société d'appartenance (facultatif)			
3 Nom			
Prénoms			
Adresse	Rue		
	Code postal et ville		
Société d'appartenance (facultatif)			
S'il y a plus de trois inventeurs, utilisez plusieurs formulaires. Indiquez en haut à droite le N° de la page suivi du nombre de pages.			
DATE ET SIGNATURE(S)			
DU (DES) DEMANDEUR(S)			
OU DU MANDATAIRE			
(Nom et qualité du signataire)		Le 25 juin 2003 CABINET PLASSERAUD Eric BURBAUD 94-0304	

La loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire. Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI.

